DOCUMENT RESUME

ED 450 155 TM 032 345

AUTHOR	George, Carrie A.
TITLE	Evaluating Assumptions of Multivariate Normality.
PUB DATE	2001-02-00
NOTE	31p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (New Orleans, LA, February 1-3, 2001).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Estimation (Mathematics); *Evaluation Methods; *Multivariate Analysis; *Statistical Distributions
IDENTIFIERS	*Normal Distribution; Normality Tests; *Univariate Analysis

ABSTRACT

Multivariate techniques have been implemented with greater and greater frequency. In order to use multivariate techniques researchers must understand the fundamental assumptions. The purpose of this paper is to evaluate one of the assumptions of multivariate analysis, normality. Overall, normal distributions are unimodal and symmetrical, and they have coefficients of skewness and kurtosis equal to 0. The specific requirements for univariate and multivariate normality are discussed. Graphical and statistical techniques are reviewed that estimate univariate, bivariate, and multivariate normality. The use of specialized computer programs for multivariate normality is also discussed. An appendix contains the Statistical Package for the Social Sciences syntax for evaluating multivariate normality. (Contains 2 tables, 7 figures, and 13 references.) (Author/SLD)

0000000284159

Reproductions supplied by EDRS are the best that can be made from the original document.

Running Head: EVALUATING ASSUMPTIONS OF MULTIVARIATE NORMALITY

ED 450 155

Evaluating Assumptions of Multivariate Normality

Carrie A. George

Texas A & M University

0000000028460

TM032345

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

C. A. George

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Paper presented at the annual meeting of the Southwest Educational Research Association, New Orleans, February 1-3, 2001.

BEST COPY AVAILABLE

Abstract

Multivariate techniques have been implemented with greater and greater frequency. In order to use multivariate techniques researchers must understand the fundamental assumptions. The purpose of the present paper is to evaluate one of the assumptions of multivariate analyses, normality. In the paper, univariate and bivariate normality will be explored. Then, graphical and statistical techniques will be reviewed to estimate univariate, bivariate and multivariate normality. Finally, the use of specialized computer programs for multivariate normality will be discussed.

000000028461

Evaluating Assumptions of Multivariate Normality

In the social sciences, researchers function with the understanding that there are usually multiple causes and multiple consequences for the effects that are found. In order to understand these complex effects, multivariate techniques are being implemented with greater and greater frequency. Multivariate analyses are critical to the understanding of the social sciences for two reasons (Fish, 1988; Thompson, 1994). First, by using multivariate techniques the researcher avoids the inflation of experimentwise Type I error rates that can occur when multiple univariate techniques are used in a single study. Second, multivariate techniques honor the reality in which researchers within the social sciences work. The world studied involves multiple variables so, in turn, multivariate techniques are essential.

To understand the increasing frequency by which multivariate analyses are being used Emmons, Stallings and Layne (1990) evaluated 16 years of research in three separate journals. They found that, "The multivariate characteristic of the social science research environment with its many confounding or intervening variables has been addressed through the trend toward increased use of multivariate analyses of variance and covariance, multiple regression, and multiple correlation" (p. 14). Grimm and Arnold (1995) also discussed that, "In the last 20 years, the use of multivariate statistics has become commonplace. Indeed it is difficult to find empirically based

TM028462

ERIC

articles that do not use one or another multivariate analysis" (p. vii).

The purpose of the present paper is to understand the assumptions of multivariate analyses and to discover the problems in violating those assumptions. In particular, the assumption of multivariate normality will be explored. According to Marascuilo and Levin (1983)

> The multivariate normal distribution is somewhat hidden throughout multivariate methods. It is not required in the estimation and data description aspects of the theory. Its impact and role, however, are basic to the [statistical significance] inference procedures of multivariate analysis and it is here that it must be assumed. There are no satisfactory tests of its truth in any one situation. (p. 203)

Multivariate normality is not required when estimating function coefficients or structure coefficients (i.e. parameter estimation), but when evaluating the results of the multivariate analyses, the underlying assumption is that the distributions are normal. In other words, in order to compare multiple variables, the data must be normally distributed.

Also, to understand the multivariate parameters being estimated researchers study the variance/covariance matrix from the sample. According to Thompson (1984), multivariate normality is critical to the interpretation of the variance/covariance matrix in that

0000028463

> the magnitudes of the coefficients of the correlation [or covariance] matrix. . . [can be] attenuated by large differences in the shapes of the distributions for the variables. It is important to emphasize that. . . [parameter estimation usually] does not require that the variables be normally distributed as long as there is no substantial attenuation associated with distribution differences, regardless of what these distributions may be. (p. 17)

In order to understand the importance of the multivariate normality assumption, univariate and bivariate normality will first be reviewed. Graphical and statistical techniques for estimating univariate, bivariate and multivariate normality will also be explored. Finally, the use of a specialized computer program (e.g., Thompson, 1990) for estimating multivariate normality will be discussed.

Assumptions of Multivariate Analyses

The first assumption of multivariate analyses is that the variance/covariance matrices are equal. This is known as homogeneity of the variance/covariance matrices. Researchers can use the Box's M, a statistical test for bivariate correlations, to determine if the variance/covariance matrices are equal. In this case, the null hypothesis is that the variance/covariance matrices are equal. You do not want to reject this null hypothesis.

The second assumption, and the focus of this paper, is that the underlying distributions are normal. As noted earlier, in

0000000028464

order to compare the variables in a multivariate analysis, the variables should all be normally distributed. The Box's _M_ test can also be implemented to test this assumption, because it is sensitive to multivariate normality. Unfortunately, if you do _not_ reject the null that the variance/covariance matrices are equal, you _might_ have a problem with multivariate normality. Understanding the assumptions of multivariate analyses is important, but what are the criteria for normal distributions? In the next two sections, the criteria for both univariate and bivariate normality will be discussed and examples will be provided.

Univariate Normality

In univariate analyses one dependent variable is being studied. As stated, most research does not involve only one dependent variable; However, understanding univariate analyses is a critical stepping stone to understanding more complex analyses.

In the univariate case, a visual inspection of the graphical data can help in the initial stages. However, visually inspecting the data is not enough. Certain statistical values are needed for the univariate distribution to be normal.

Skewness

Investigating the skewness of a distribution is part of the determination of univariate normality. Skewness is a measure of the symmetrical shape of the distribution. The normal distribution has a skewness value of 0, which means it is perfectly symmetrical. Being symmetrical does not necessarily mean "bell shaped" as bimodal and rectangular/uniform

0000000028465

distributions may also have a skewness value of 0. Any distribution can be positively (skewed right) or negatively skewed (skewed left). The tail of a positively skewed distribution extends to the right and the tail of a negatively skewed distribution extends to the left.

Kurtosis

Another factor to investigate is the kurtosis of the distribution. Kurtosis is also a measure of the shape of the distribution. It is a measure of the height relative to the width. Distributions that are high and narrow are leptokurtic and they will have positive kurtosis values. Distributions that are low and wide are platykurtic (much like a platypus) and will have negative kurtosis values. Distributions, like those that resemble the "bell curve", can have mesokurtic distributions, which are closer to the value of 0. It is interesting to note that platykurtic univariate distributions can be a sign of power problems in the data analyses, because a majority of the scores are falling into the tails of the distribution.

General Concepts

Overall, normal distributions are unimodal, symmetrical and they have coefficients of skewness and kurtosis equal to 0. According to Ferguson (1976), the more specific criteria for univariate normal distribution are (that)

1. The curve is symmetrical. The mean, median, and mode coincide.

0000028466

2. The maximum ordinate of the curve occurs at the mean, that is, where *z*=0 and in the unit normal curve is equal to .3989

3. The curve is asymptotic. It approaches but does not meet the horizontal axis and extends from minus infinity to plus infinity.

4. The points of inflection of the curve occur at points +/- 1 standard deviation unit above and below the mean. Thus, the curve changes from convex to concave in relation to the horizontal axis at these points.

5. Roughly 68 percent of the area of the curve falls within the limits +/- 1 standard deviation unit from the mean.

6. In the unit normal curve the limits *z*= +/- 1.96 include 95 percent and the limits *z*= +/- 2.58 include 99 percent of the total area of the curve, 5 percent and 1 percent of the area, respectively, falling beyond these limits. (pp. 93-94)

Graphical and Statistical Techniques for Determining Univariate Normality

The researcher should first evaluate the shape of the distribution visually (however, visual inspection is not entirely sufficient). Computer software, such as SPSS, is useful for this evaluation. Statistical significance testing can also be implemented. The steps for determining univariate normality in SPSS are to

(a) Have the computer program plot the data with a histogram and then select the options that allow the normal

0000002846/

ERIC

curve to be drawn over the histogram. Visually inspect the results.

INSERT FIGURE 1 ABOUT HERE

(b) Use other graphical techniques like Q-Q plots, box-plots, and stem-and-leaf plots to visually inspect the data.

(c) Use the Explore option to evaluate the skewness and kurtosis values of the data in the distribution.

(d) Use the Kolmogorov-Smirnoff or Shapiro-Wilk statistic to determine if the null hypothesis (that the distribution is normal) should be rejected.

Q-Q plots, box plots, and stem-and-leaf plots. Q-Q plots (or quantile vs. quantile plots) are useful in graphically inspecting the data. According to Stevens (1996), these plots are very popular in evaluating univariate normality. Q-Q plots are created through several steps. First, the scores in the data are ranked from lowest to highest. Second, the actual scores are converted to *z*-scores and compared with expected normal values. Burdenski (2000) states that "the expected normal value is the *z*-score that a case with that rank holds in the normal distribution" (p. 10). Third, a bivariate scatterplot is created that graphically compares the actual scores and the expected scores. If the distribution is normal, then a straight diagonal line running from the lower left corner to the upper right corner will appear. Some scores will not fall along the straight line

ERIC

which can indicate outliers and possible deviation from normality.

INSERT FIGURE 2 ABOUT HERE

Box-plots, also referred to as box-and-whisker plots, are another graphical technique for evaluating normality. Box-plots allow for a bird's eye view of the data. The main body of scores will fall within the box that contains the median score. The whiskers, or vertical lines that extend outside the box, will indicate the 25th and 75th percentile. Outliers are represented as dots outside range of the whiskers. In a normal distribution the median should be in the middle of the graphical box. In addition, outliers should not appear. Deviations from these conditions can indicate problems with univariate normality.

INSERT FIGURE 3 ABOUT HERE

Stem-and-leaf plots are yet another way to evaluate univariate normality. Stem-and-leaf plots are basically histograms that are turned on their sides. In the stem-and-leaf plot, "data values are collected into intervals and displayed as bars. . .[in which] the digits of each number are separated into a stem and a leaf" (SPSS Base 9.0 Applications Guide, 1999). The stem can be represented as the "tens" digit or another other suitable number, and the leaf can represent "ones" or "units" digit. Stem-and-leaf graphs "provide the data analyst with a

0000000028469

quick way to illustrate a distribution of scores while maintaining the actual scores in the displays" (Hinkle, Wiersma, & Jurs, 1998, p. 26). For example, if data were collected on the age of participants, the stem could be 2 for 20-year old participants and the leaf could be 2,2,3,3,5,6,9 for 22, 22, 23, 23, 25, 26, and 29-year old participants, respectively. A "bell shaped" curve (turned on its side) should appear when visually inspecting the stem-and-leaf plot. Again, deviations from this shape could indicate problems with normality.

INSERT FIGURE 4 ABOUT HERE

Skewness, kurtosis, and statistical signigicance. To evaluate the skewness and kurtosis of the distribution, utilize the Explore option in SPSS. The Explore option will provide the researcher with information concerning the mean, the standard deviation, and the skewness and kurtosis of the data. Skewness and kurtosis values that differ from 0 indicate non-normal distributions. It is critical to review these values because a visual inspection of the graphs is not enough to determine univariate normality.

INSERT TABLE 1 ABOUT HERE

SPSS also offers the researcher the opportunity to determine if the univariate distribution is normal through statistical

0000000028470

significance testing. For data sets with less than 50 cases the Shapiro-Wilk statistic can be used, and for data sets that have greater than 50 cases, the Kolmogorov-Smirnov statistic (with the Lillefors correction) can be used. The Lillefors correction, part of the Kolmogorov-Smirnov statistic, is applied when the mean and variance of the true population are not known (SPSS Base 9.0 Applications Guide, 1999). The null hypothesis being tested is that the distribution is normal. In this situation, the researcher does not want to reject the null.

INSERT TABLE 2 ABOUT HERE

Bivariate Normal Distributions

For bivariate distributions "univariate normality is a necessary but not sufficient requirement for bivariate normality" (Henson, 1999, p. 195). In bivariate distributions, two sets of data are being compared/correlated. The two sets of data, or the two distributions, should be normally distributed before the comparison is made. If they are normal, then it is possible that the bivariate distribution will be normal. According to Burdenski (2000), the characteristics of bivariate normal distributions are that

1. For each value of X, the distribution of it's associated Y values is a normal distribution and vice-versa.
2. The Y means for each value of X are linear (i.e. they fall on a straight line) and the same is true for the X means for each value of Y.

ERIC

0000002847/1

3. The scatter plots demonstrate homoscedasticity--the variance in the Y values is uniform across all values of X and the variance in the X values is constant for all values of Y. (p. 17)

Visualizing the bivariate distribution requires knowledge of what is occurring. Instead of the two-dimensional space that is needed for univariate distributions, three-dimensional space is now required. There is a set of scores (or a distribution) represented by the X-axis, a set of scores (or a distribution) represented by the Y-axis, and a frequency count of scores represented on a third axis. The three-dimensional graph that will emerge in a bivariate normal distribution will resemble a "hat" shape. If one could take a knife and slice through the bivariate normal distribution (or "hat"), the slices would appear univariate normal.

INSERT FIGURE 5 ABOUT HERE

Much like turning a bowl upside down and drawing a circle around it, there will be a two-dimensional footprint left to represent the three-dimensional bivariate distribution. Figure 5 is a two-dimensional image of a bivariate distribution. The two-dimensional image can also be compared to a bird's eye view, looking down on the bivariate distribution from above.

For the bivariate distribution, there will be a plotted mean for both of the univariate distributions known as the centroid. Concentric circles, known as contour lines, can be drawn outside

ERIC

of the centroid to represent distances of 1-3 standard deviations from the central mean. The standard deviations of both variables are used to form the contour lines. If the standard deviations of both variables are equal, then the contour lines will be circular. If the standard deviations of both variables are unequal, then the contour lines will be elliptical (Henson, 1999).

Like in the univariate case, approximately 68% of the scores should fall within 1 standard deviation of the centroid, 95% of the score should fall within 2 standard deviation of the centroid, and 99% of the scores should fall within 3 standard deviations of the centroid. If these values do not represent the actual scores, then there is a possibility that the bivariate distribution is not normal.

INSERT FIGURE 6 ABOUT HERE

In order to understand the concept that univariate normality is necessary, but not sufficient for bivariate normality, Henson (1999) discussed two data sets that were essentially univariate normal (with skewness and kurtosis values near 0), but not bivariate normal. In one pairwise combination of the data, the three-dimensional figure that appeared was not "hat" shaped, but rectangular in shape. The univariate variables that were combined did not create the typical shape of a bivariate normal distribution, thus indicating problems with bivariate normality. For a further discussion, please see Henson (1999).

ERIC

0000002847

Graphical and Statistical Techniques for Evaluating Bivariate Normality

Graphical techniques include scatterplots. In reviewing the scatterplot, the centroid should be located and contour lines should be drawn. Determine if approximately 99% of the scores fall within three-standard deviations of the centroid. If data can be plotted in three-dimensional space, then look for the "hat" shape described previously. Images that do not conform reveal non-normal bivariate distributions.

Unlike the univariate case, SPSS does not appear to have statistical significance tests for bivariate normality. It is possible to use the Kolmogorov-Smirnov or Shapiro-Wilk statistic to determine if the univariate distributions are normal and then to apply the information to the bivariate distribution. Keep in mind, however, that univariate statistical significance will not be sufficient for bivariate statistical significance.

Multivariate Normality

In multivariate analyses, two or more dependent variables are being studied. All of the variables must be univariate normal and "all possible pairs of the variables must be normal" (Burdenski, 2000, p. 19). However, proving bivariate normality will not suffice, because "bivariate normality is a necessary, but not sufficient" (Henson, 1999, p. 195) requirement for multivariate normality. Unlike the univariate and bivariate situations, SPSS does not currently offer statistical or graphical tests for multivariate normality. The question remains about how to determine multivariate normality.

To answer this question, a discussion of the Malhalanobis distance ($\underline{D}^2$) is necessary. The Malhalanobis distance is the "distance of a case from the centroid [or the mean vector of scores] where the centroid [mean vector] is the point defined by the means of all the variables taken as a whole" (Burdenski, 2000). Observations far from the centroid are possible outliers that may contribute to non-normality. The Malhalanobis distance is a favorable measure of normality because it is independent of sample size. This makes the $\underline{D}^2$ value superior to the statistical significance tests previously mentioned. The Malhalanobis distance is understood in terms of the following formula:

$$D_i^2=(x_i-M)'S-1(x_i-M),$$

where $\underline{D}_i^2$ is the Malhalanobis distance for an individual, and $\underline{S}$ is the variance/covariance matrix (Burdenski, 2000), $\underline{x}_i$ is the "vector of the data for case $\underline{i}$ and $\underline{M}$ is the vector of means (centroid) for the predictors" (Stevens, 1996, p. 111).

According to Thompson (1990), a statistical program named MULTINOR can be used to calculate multivariate normality). MULTINOR utilizes SPSS syntax to implement the program (see Appendix). Through MULTINOR, the $\underline{D}^2$ value for each observation is calculated in a standardized format taking into account variability between each variable and the correlations that exist between the variables (Henson, 1999). A graphical representation of the $\underline{D}^2$ values can be plotted on a scatterplot. If the dependent variables (in combination) are multivariate normal, then the plot will form a straight diagonal line that runs from

ERIC

000000028475

the lower left corner to the upper right corner of the graph. This concept is similar to the Q-Q plots for the univariate case.

INSERT FIGURE 7 ABOUT HERE

When viewing the scatterplot look for outliers that can attenuate multivariate normality. Also determine if the line fits the pattern for normality. Evaluating the line is solely up to researcher judgment, as no formal standards are in place.

Conclusion

While multivariate normality is not needed to calculate function or structure coefficients, it is important to evaluate. Normality is a critical assumption of multivariate analyses that should not be overlooked. Ignoring normality issues in data sets can lead to misinformed results. The researcher should analyze normality in every situation, because decisions based on non-normal data sets will be faulty.

0000002847G

ERIC

References

Burdenski, T. (2000, April). Evaluating univariate, bivariate, and multivariate normality using graphical procedures. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Emmons, N. J., Stallings, W. M., & Layne, B. H. (1990, April). Statistical methods used in *American Educational Research Journal, Journal of Educational Psychology,* and *Sociology of Education* from 1972 through 1987. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA. (ERIC Document Reproduction Service No. ED 319 797)

Ferguson, G. A. (1976). Statistical analysis in psychology and education (4th ed.). New York: McGraw Hill.

Fish, L. (1988). Why multivariate methods are usually vital. Measurement and Evaluation in Counseling and Development, 21, 130-137.

Grimm, L. G., & Yarnold, P. R. (Eds.). (1995). Reading and understanding multivariate statistics. Washington, DC: American Psychological Association.

Henson, R. K. (1999). Multivariate normality: What is it and how is it assessed? In B. Thompson (Ed.), Advances in social science methodology (Vol. 5, pp. 193-212). Stamford, CT: JAI Press.

Hinkle, D.E., Wiersma, W., & Jurs, S. G. (1998). Applied statistics for the behavioral sciences (4th ed.). Boston: Houghton Mifflin Company.

0000028477

Marascuilo, L. A., & Levin, J. R. (1983). Multivariate statistics in the social sciences: A researcher's guide. Monterey, CA: Brooks/Cole.

SPSS Base 9.0 Applications Guide. (1999). Chicago: Author.

Stevens, J. (1996). Applied multivariate statistics for the social sciences (3rd ed.). Mahaway, NJ: Lawrence Erlbaum Associates.

Thompson, B. (1984). Canonical correlation analysis: Uses and interpretation. Thousand Oaks, CA: SAGE.

Thompson, B. (1990). MULTINOR: A FORTRAN program that assists in evaluating multivariate normality. Educational and Psychological Measurement, 50, 845-848.

Thompson, B. (1994, February). Why multivariate methods are usually vital in research: Some basic concepts. Paper presented as a Featured Speaker at the biennial meeting of the Southwestern Society for Research in Human Development, Austin, TX. (ERIC Document Reproduction Service No. ED 367 687)

0000028478

ERIC

Appendix

SPSS Syntax for Evaluating Multivariate Normality

```
COMMENT 'y' is a variable automatically created by the program,
COMMENT it does not have to be modified for different data sets.
COMPUTE y=$casenum .
PRINT FORMATS y(F5) .
REGRESSION
  /DESCRIPTIVES MEAN STDDEV CORR SIG N
  /MISSING LISTWISE
  /DEPENDENT y
  /METHOD=ENTER t1 t3 t13
  /SAVE MAHAL .
SORT CASES BY MAH_1(A) .
EXECUTE .

LIST VARIABLES=y t1 t3 t13 MAH_1
  /FORMAT=NUMBERED .
COMMENT In the next TWO command lines, for a given data set put
the
COMMENT actual n in place of the number '301' used for this
example.
LOOP #i=1 to 301 .
COMPUTE p=($casenum - .5) / 301 .
COMMENT In the next line, change '3' to whatever is the number
COMMENT of variables for which you are assembling normality
COMMENT The p critical value of chi square for a given case
COMMENT is set as {[the case number (after sorting) - .5] / the
COMMENT sample size}.
COMPUTE chisq=idf.chisq (p,3) .
END LOOP .
PRINT FORMATS p chisq (F8.5) .
LIST VARIABLES=y p MAH_1 chisq
  /FORMAT=NUMBERED .
PLOT
  VERTICAL='CHI SQUARE'/
  HORIZONTAL='MALHALANOBIS DISTANCE'/
  PLOT=chisq with MAH_1 .
```

Table 1

Means, Standard Deviations, Skewness and Kurtosis in Explore

Descriptives

			Statistic	Std. Error
T1	Mean		29.61	.40
	95% Confidence Interval for Mean	Lower Bound	28.82	
		Upper Bound	30.41	
	5% Trimmed Mean		29.70	
	Median		30.00	
	Variance		49.064	
	Std. Deviation		7.00	
	Minimum		4	
	Maximum		51	
	Range		47	
	Interquartile Range		9.00	
	Skewness		-.257	.140
	Kurtosis		.355	.280

Descriptives

			Statistic	Std. Error
T3	Mean		14.23	.16
	95% Confidence Interval for Mean	Lower Bound	13.91	
		Upper Bound	14.55	
	5% Trimmed Mean		14.21	
	Median		14.00	
	Variance		8.011	
	Std. Deviation		2.83	
	Minimum		6	
	Maximum		25	
	Range		19	
	Interquartile Range		4.00	
	Skewness		.206	.140
	Kurtosis		.722	.280

0000000028480

ERIC Full Text Provided by ERIC

Table 2

Kolmogorov-Smirnov Tests of Statistical Significance for Univariate Distributions

Tests of Normality

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
T1	.060	301	.011

a. Lilliefors Significance Correction

Tests of Normality

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
T3	.094	301	.000

a. Lilliefors Significance Correction

ERIC

Figure Captions

Figure 1. Histogram with normal curve example

Figure 2. Q-Q plot example

Figure 3. Box-plot example

Figure 4. Stem-and-leaf plot example

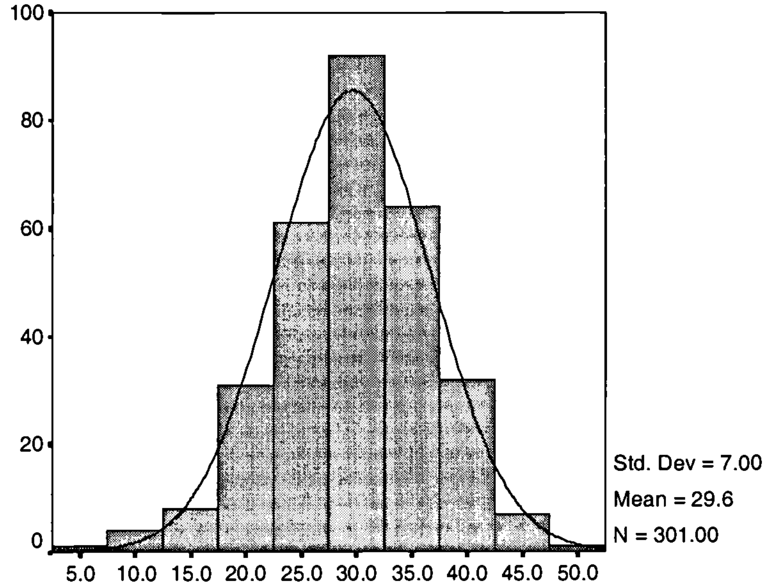
Figure 5. Three-dimensional "hat" formation of bivariate data

Figure 6. Bivariate contour lines representing 1 standard deviation, 2 standard deviations, and 3 standard deviations away from the centroid.

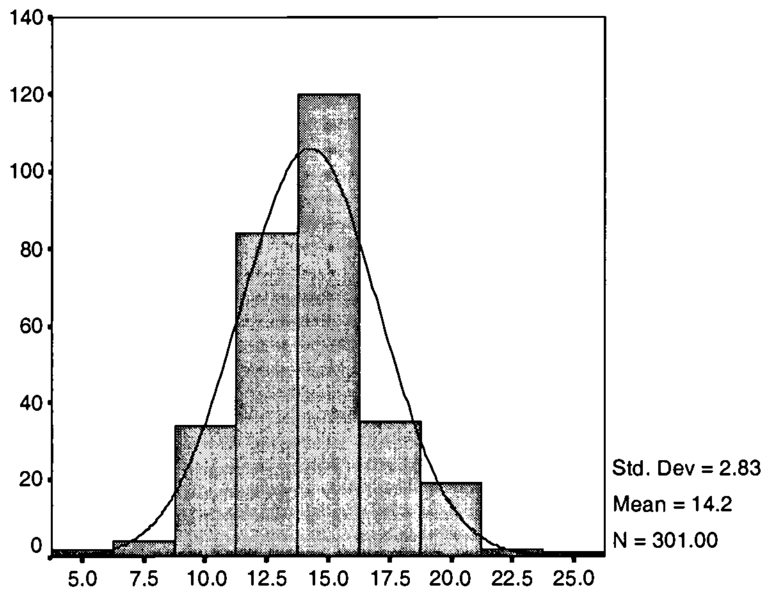
Figure 7. Multivariate normality scatterplot

0000028482

Figure 1.

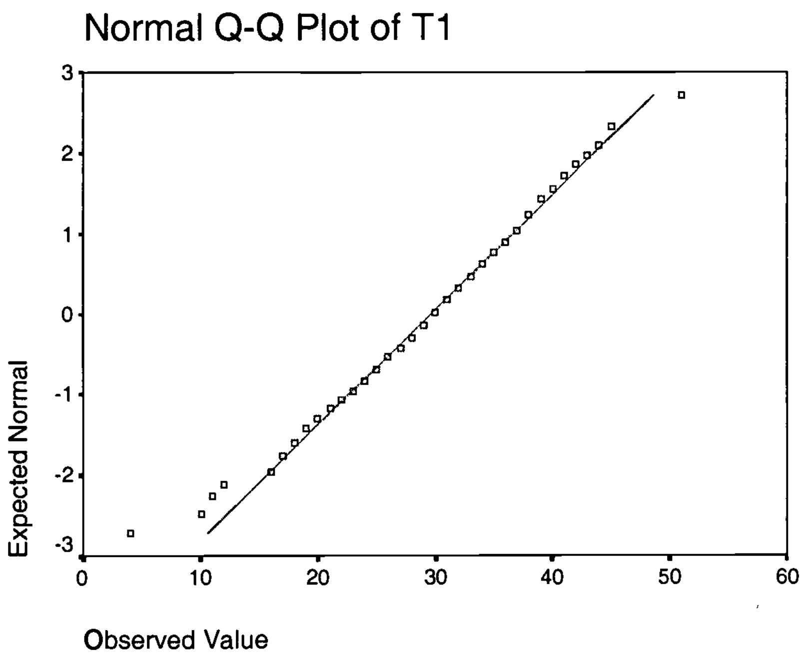


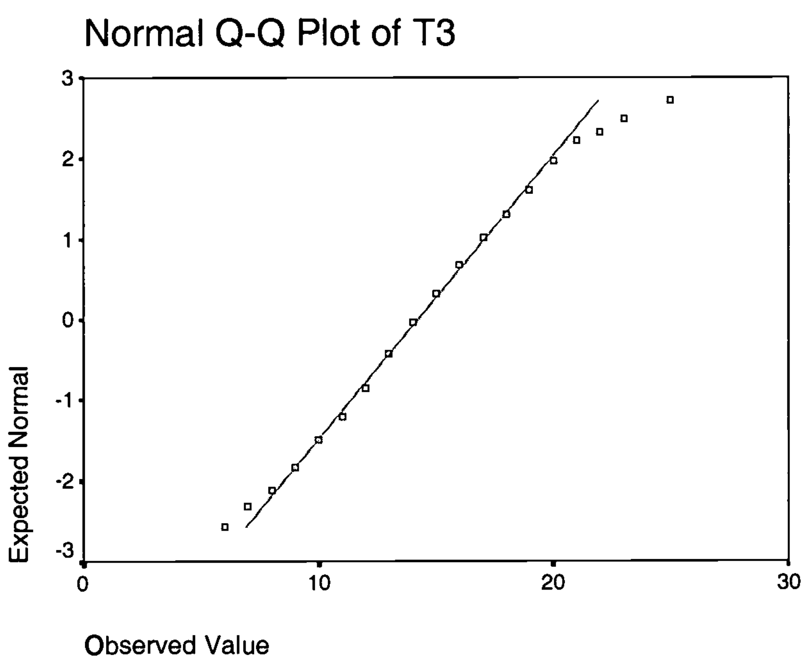
T1



T3

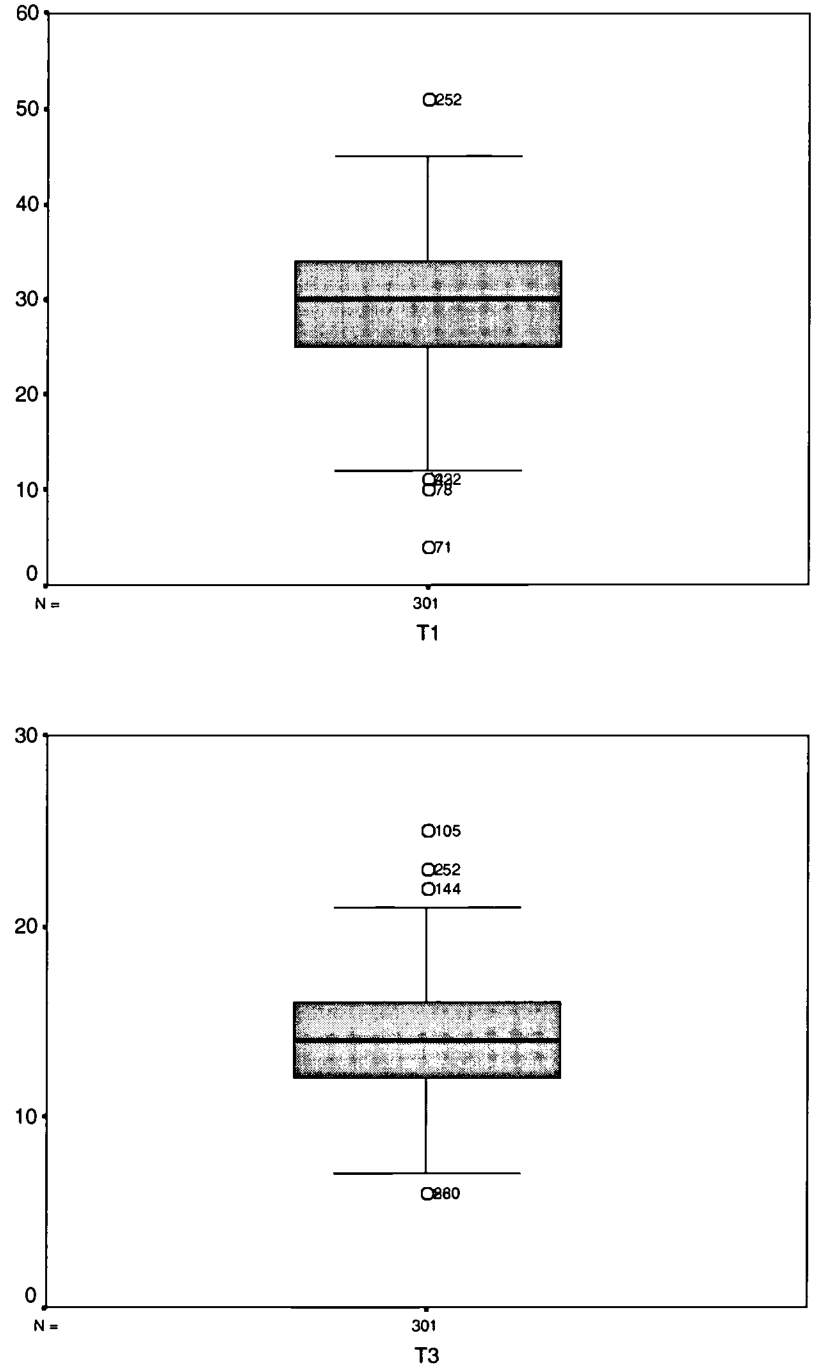
Figure 2.





0000028484

Figure 3.



ERIC

Figure 4.

T1 Stem-and-Leaf Plot

```
 Frequency    Stem &  Leaf

     4.00 Extremes    (=<11)
      .00        1 .
     1.00        1 .  2
      .00        1 .
     8.00        1 .  66667777
    13.00        1 .  8888889999999
    14.00        2 .  00000011111111
    15.00        2 .  222233333333333
    27.00        2 .  444444444455555555555555555
    23.00        2 .  66666666666667777777777
    37.00        2 .  8888888888888888888899999999999999999
    37.00        3 .  0000000000000000000000111111111111111
    33.00        3 .  222222222222222222333333333333333
    29.00        3 .  44444444444444444555555555555
    20.00        3 .  66666666667777777777
    20.00        3 .  88888888888888889999
    10.00        4 .  0000011111
     4.00        4 .  2233
     5.00        4 .  44555
     1.00 Extremes    (>=51)

 Stem width:  10
 Each leaf:       1 case(s)
```

T3 Stem-and-Leaf Plot

```
 Frequency    Stem &  Leaf

     2.00 Extremes    (=<6.0)
     1.00        7 .  0
     3.00        8 .  000
     7.00        9 .  0000000
    13.00       10 .  0000000000000
    14.00       11 .  00000000000000
    38.00       12 .  00000000000000000000000000000000000000
    46.00       13 .  0000000000000000000000000000000000000000000000
    45.00       14 .  000000000000000000000000000000000000000000000
    39.00       15 .  000000000000000000000000000000000000000
    36.00       16 .  000000000000000000000000000000000000
    21.00       17 .  000000000000000000000
    14.00       18 .  00000000000000
    12.00       19 .  000000000000
     6.00       20 .  000000
     1.00       21 .  0
     3.00 Extremes    (>=22)

 Stem width:   1
 Each leaf:       1 case(s)
```

Figure 5.

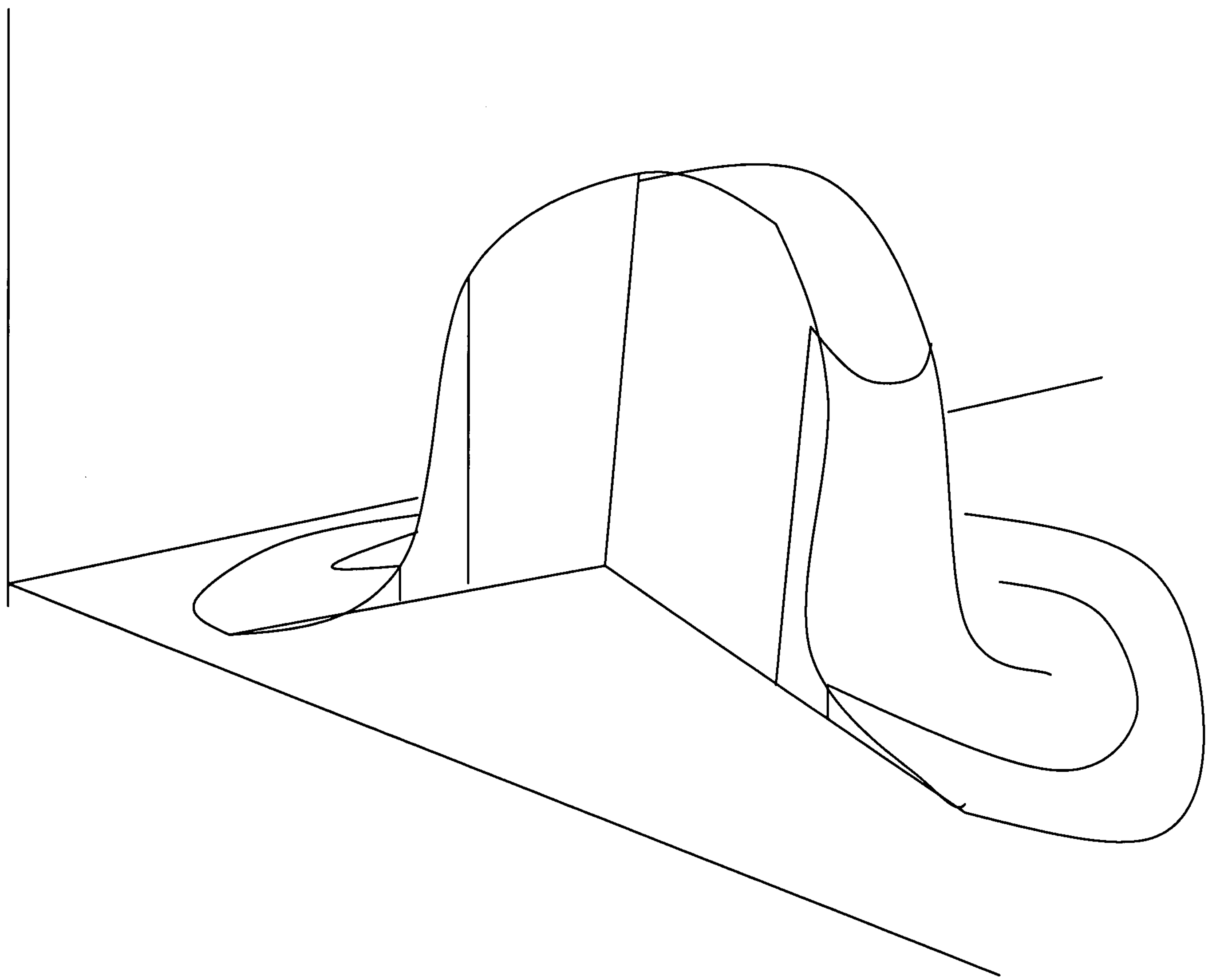
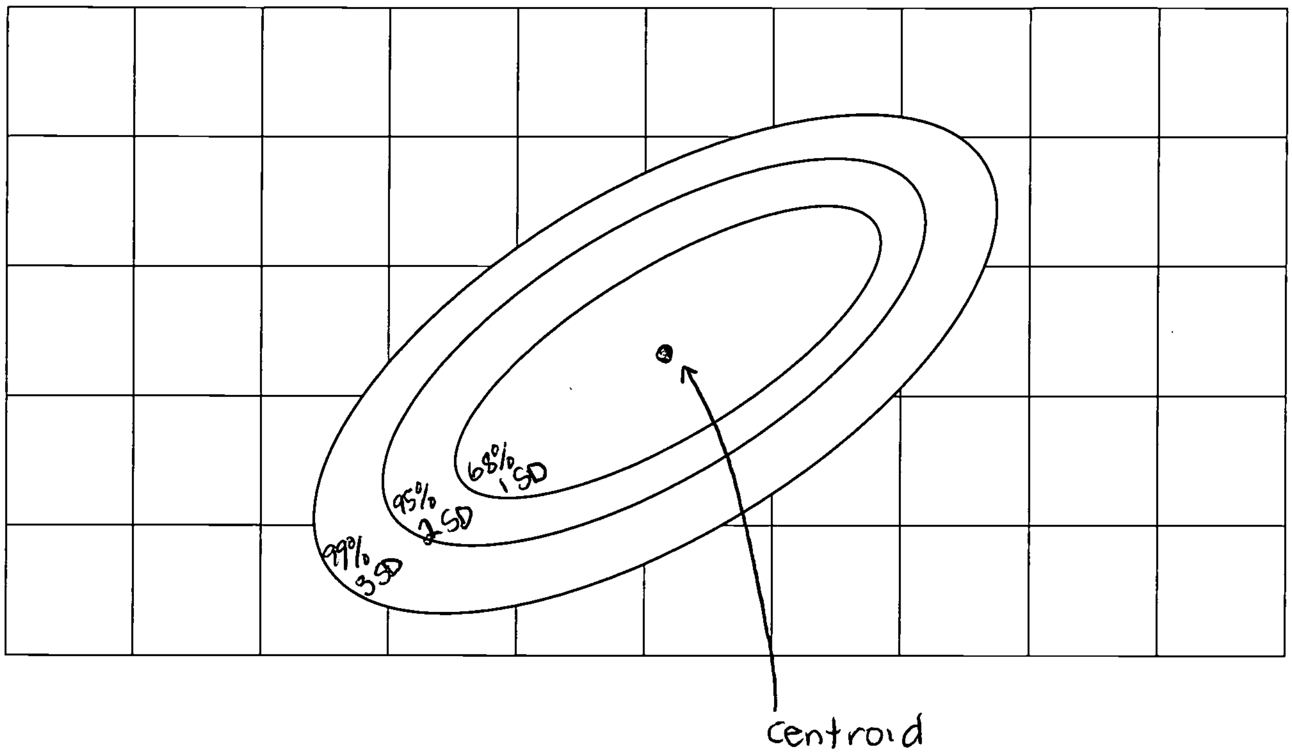


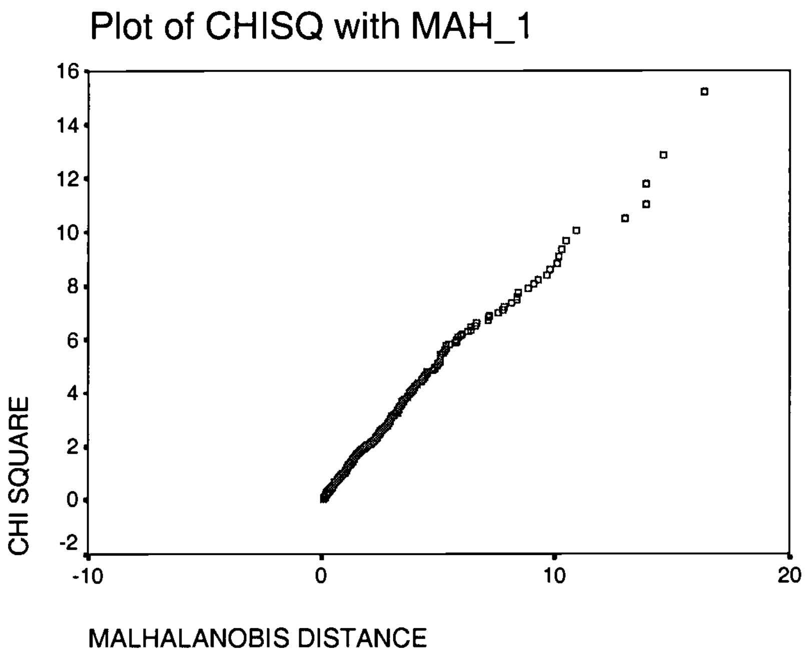
Figure 6.



$\underline{r}$=.7

BEST COPY AVAILABLE

Figure 7.



ERIC

TM032345

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE
(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: EVALUATING ASSUMPTIONS OF MULTIVARIATE NORMALITY

Author(s): CARRIE A. GEORGE

Corporate Source:

Publication Date: 2/1/01

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

[X] Sample sticker to be affixed to document

Check here
Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
CARRIE A. GEORGE
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document [ ]

or here
Permitting reproduction in other than paper copy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY
Sample
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: X Carrie George	Position: RES ASSOCIATE
Printed Name: CARRIE A. GEORGE	Organization: TEXAS A&M UNIVERSITY
Address: TAMU DEPT EDUC PSYC COLLEGE STATION, TX 77843-4225	Telephone Number: 979/845-1335
	Date: 1/22/01

OVER